



DATOS IDENTIFICATIVOS

Análise Exploratoria de Datos

Materia	Análise Exploratoria de Datos			
Código	V03M100V01104			
Titulación	Máster Universitario en Técnicas Estadísticas			
Descritores	Creditos ECTS	Sinale	Curso	Cuadrimestre
	6	OB	1	1c
Lingua de impartición				
Departamento				
Coordinador/a	Roca Pardiñas, Javier			
Profesorado	Iglesias Pérez, María Carmen Roca Pardiñas, Javier			
Correo-e	roca@uvigo.es			
Web	http://eio.usc.es/pub/mte/			
Descrición xeral	Profesorado: Javier Roca Pardiñas (UVigo): 1.5 ECTS Ana Pérez González (UVigo): 3 ECTS M. Carmen Iglesias Pérez (UVigo): 1.5 ECTS			
	Más información en http://eio.usc.es/pub/mte/			

Competencias de titulación

Código	
A1	Capacidad para comprender, presentar, formular y resolver aquellos problemas susceptibles de ser abordados a través de modelos de la estadística
A3	Conocer las aplicaciones de los modelos de la estadística
A5	Coñecer algoritmos de resolución de los problemas y manejar el software adecuado

Competencias de materia

Resultados previstos na materia	Resultados de Formación e Aprendizaxe
Comprender la finalidad, resultados y beneficios del análisis de un conjunto de datos, así como de sus requerimientos, con el fin de permitir una mejor modelización de problemas y experimentos.	A1 A3
Descubrir la problemática del análisis de un conjunto de datos, con el fin de motivar conceptos y técnicas de la teoría de la probabilidad y de la estadística matemática.	A1 A3 A5
Conocer los conceptos, y técnicas numéricas y gráficas, esenciales para el descubrimiento y entendimiento de las estructuras y relaciones contenidas en un conjunto de datos.	A1 A3
Adquirir una visión general de algunos de los principales problemas estadísticos y de los grupos de técnicas más apropiadas para resolverlos.	A1 A3 A5
Adquirir una formación básica en el manejo de herramientas estadísticas a través del programa R (http://www.r-project.org).	A5

Contidos

Tema	
Manejo del software estadístico R	Descarga e instalación. Estructuras de datos. Lectura, importación y exportación de datos. Tratamiento de datos. Procedimientos gráficos. Introducción a la programación.

Análisis descriptivo unidimensional	Población, muestra, tipos de variables. Medidas descriptivas numéricas, creación de gráficos, detección de datos atípicos, estudio de la normalidad de los datos. Principales distribuciones discretas (binomial, Poisson, etc.) y continuas (gaussiana, exponencial, etc.).
Análisis de variables cualitativas	Tablas de contingencia y medidas de asociación. Análisis de variables numéricas: gráficos y medidas de correlación. Comparación de distribuciones: test paramétricos y no paramétricos.
Análisis de la varianza y modelos de regresión	Análisis de la varianza. Modelos de regresión paramétricos y no paramétricos.
Introducción al análisis de datos multivariantes	análisis de componentes principales, análisis factorial, análisis de conglomerados y análisis discriminante.

Planificación

	Horas na aula	Horas fóra da aula	Horas totais
Prácticas de laboratorio	12	30	42
Tutoría en grupo	0	20	20
Sesión maxistral	22	38	60
Resolución de problemas e/ou exercicios	10	10	20
Probas prácticas, de execución de tarefas reais e/ou simuladas.	4	4	8

*Os datos que aparecen na táboa de planificación son de carácter orientador, considerando a heteroxeneidade do alumnado.

Metodoloxía docente

	Descrición
Prácticas de laboratorio	La docencia se desenvolverá mediante la resolución de problemas reais o simulados utilizando los modelos tratados en las sesiones magistrales. Se utilizará principalmente el software R.
Tutoría en grupo	Se mantendrá un servizo de tutoría en grupo a los alumnos. Los alumnos tamén podrán consultar sus dúbidas por correo electrónico.
Sesión maxistral	La docencia se desenvolverá mediante la exposición por parte del profesor de las diferentes técnicas de Análisis Exploratorio de Datos Para ello, los alumnos dispondrán de apuntes elaborados que servirán de material básico para el estudio y en su defecto de material e información sobre bibliografía específica disponible en la biblioteca o en internet.

Atención personalizada

Metodoloxías	Descrición
Tutoría en grupo	La tutoría se desenvolverá intercalada entre las sesiones magistrales y las prácticas de laboratorio con el fin adicional de evaluar el grado de asimilación de la materia por parte de los alumnos, así como mediante correo electrónico individual o mediante un foro en el que se publicarán las preguntas formuladas y las respostas proporcionadas por el profesor.

Avaliación

	Descrición	Cualificación
Resolución de problemas e/ou exercicios	Realización de cuestionarios e actividades que sexan plantexadas durante 01 periodo de docencia da materia.	40
Probas prácticas, de execución de tarefas reais e/ou simuladas.	Realización dun exame final teórico práctico.	60

Outros comentarios sobre a Avaliación

Bibliografía. Fontes de información

1. Abraira Santos, V. y Pérez de Vargas, A. (1996). Métodos Multivariantes en Bioestadística. Centro de Estudios Ramón Areces.
2. Agresti, A. (1996). An Introduction to Categorical Data Analysis, John Wiley, Nueva York.
3. Crawley, M. J. (2005). Statistics: an introduction using R. Ed. John Wiley and Sons.
4. Everitt, B. S. (2005). An R and S-PLUS Companion to Multivariate Analysis. Ed. Springer
5. Everitt, B. S. (2011) An Introduction to applied multivariate analysis with R. Springer.
6. Hair J. F., Anderson, Tatham and Black (1996). Análisis multivariante. 5ª Edición. Prentice Hall.
7. Maindonald, J. H. (2007). Data analysis and graphics using R: an example-based approach. Cambridge University Press.

8. Martínez Almécija, Alfredo y otros (1993). Inferencia Estadística. Un enfoque clásico. Pirámide.
9. Peña Sánchez de Rivera, D. (1999). Estadística. Modelos y Métodos. 2. Modelos Lineales y Series Temporales. Alianza Universidad Textos.
10. Peña Sánchez de Rivera, D. (2002). Análisis de datos multivariantes. McGraw Hill, Madrid.
11. Sheather, S. J. (2009). A modern approach to regression with R. New York; London: Springer.
12. Wand, M. P. and Jones, M. C. (1995). Kernel Smoothing. Chapman and Hall.
13. Tukey, J.W. (1977). Exploratory Data Analysis. Addison-Wesley.
14. Zuur, Alain F.(2009) A Beginner's guide to R. New York . Springer.

Recomendacións

Outros comentarios

No se necesita haber cursado ninguna otra asignatura del máster. Sin embargo es fundamental la asistencia regular a las clases para la superación de esta materia, ya que es muy importante el seguimiento del trabajo realizado en el aula.

Los requisitos básicos de esta materia son un conocimiento básico de la Estadística y conocimientos a nivel usuario de Windows. Como ya se ha comentado se utilizará el software libre R.

En algunas sesiones se realizarán cuestionarios y ejercicios sobre temas impartidos para la valoración de la evolución y comprensión de los alumnos sobre la materia.
